View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Publicações BAD (Bibliotecários, Arquivistas e Documentalistas - E-Journals)

Da "The European Library" à "Europeana" – Um percurso com impulsos nacionais

José Borbinha
Instituto Superior Técnico / INESC-ID
Rua Alves Redol 9 - Apartado 13069
1000-029 Lisboa
E-mail: jlb@ist.utl.pt

Nuno Freire
Biblioteca Nacional de Portugal
Campo Grande, 83
1749-081 Lisboa
E-mail: nuno.freire@bnportugal.pt

RESUMO
Este artigo descreve o percurso a nível Europeu que levou à iniciativa TEL e desta à realidade emergente da Europeana. Neste percurso, equipas de engenharia do Instituto Superior Técnico e do INESC-ID aliadas à competência da Biblioteca Nacional de Portugal têm desempenhado um papel de grande relevo que nem sempre tem sido público. Faz-se assim também uma descrição desse envolvimento, numa perspectiva principalmente tecnológica.

PALAVRAS-CHAVE: Bibliotecas Digitais. Metadados. Arquitecturas de Sistemas. Interoperabilidade. TEL. Europeana.

INTRODUÇÃO

A iniciativa TEL – *The European Library* marcou o primeiro passo concreto de colaboração entre as bibliotecas nacionais Europeias visando o desenvolvimento e manutenção de serviços de informação efectivos (isto é, indo para além das fases típicas de projecto a que iniciativas anteriores se tinham limitado). Na sequência deste primeiro exemplo, um outro objectivo, bem mais ambicioso, ganhou forma visando criar uma verdadeira "biblioteca digital Europeia": a iniciativa Europeana – *European Digital Library*.

Neste já relativamente longo percurso temos encontrado com frequência equipas de engenharia do IST - Instituto Superior Técnico e do INESC-ID, assim como variados envolvimentos da BNP - Biblioteca Nacional de Portugal.

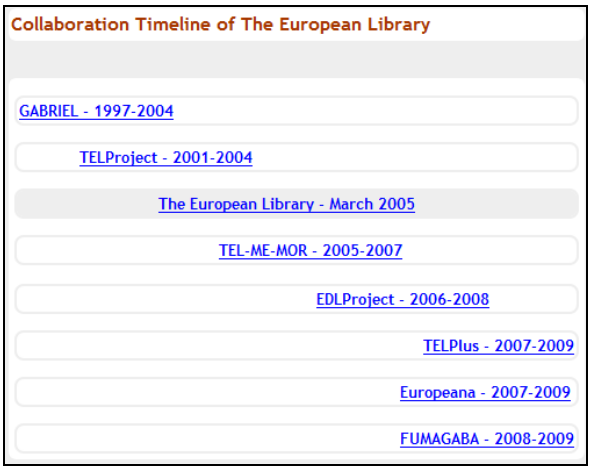

Figura 1: Projectos relevantes até 2009 [1].

A nossa história pode-se dividir duas linhas de acção (Figura 1): a linha inicial, iniciada em 1997 e que levou ao portal TEL; e a linha que arrancou em 2006 e levou à conceptualização e demonstração do conceito da Europeana. Ambas estas linhas prosseguem neste momento em paralelo, trocando lições mas mantendo objectivos diferentes, ainda que complementares. Este artigo oferece uma descrição geral deste contexto, começando por dar uma descrição sucinta de cada um dos projectos relevantes e do papel que nele estas instituições desempenharam (ou desempenham ainda na data de escrita deste artigo). Para além dos aspectos meramente descritivos será dado realce à importância da contribuição destes parceiros nacionais.

Na segunda parte deste artigo iremos apresentar uma síntese dessas realizações, onde iremos sublinhar em concreto as especializações tecnológicas.

O PORTAL GABRIEL

O primeiro caso concreto de colaboração neste domínio entre as bibliotecas nacionais Europeias deu-se com o portal Gabriel - GAteway and BRIdge to Europe's National Libraries. Esta iniciativa foi promovida pelo CENL – Conference of the European National Librarians, na qual estão representados todos os países do Conselho da Europa.

Este portal esteve em linha durante 8 anos (até Junho de 2005), sendo mantido em Inglês, Francês e Alemão. Oferecia um directório de 43 bibliotecas de 41 países (a BNP representava Portugal; Itália e Rússia eram representados por duas bibliotecas), com informação sobre organização e estrutura de cada uma, história, colecções, contactos dos serviços, e hiperligações para os sítios Web e também para as colecções e catálogos em linha.

O PORTAL TEL

O portal TEL (Figura 2 [2]) nasceu com a visão de oferecer um serviço de pesquisa simultânea em todos os catálogos em linha das bibliotecas já representadas no GABRIEL. Entre Janeiro de 2001 e Janeiro de 2004 decorreu um projecto financiado pelo programa IST – *Information Society and Technologies* da Comissão Europeia (parte do 5º Programa Quadro).

A BNP foi parceira neste projecto, em que contou com o apoio técnico do INESC-ID (que dessa forma envolveu alunos de informática do IST – Instituto Superior Técnico).

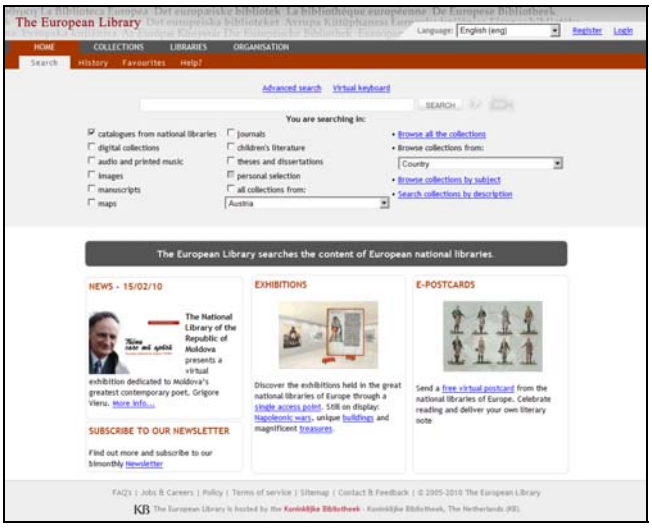

Figura 2: A interface do portal TEL.

Os restantes parceiros deste projecto pioneiro foram as bibliotecas nacionais da Finlândia, Alemanha, Holanda, Eslovénia, Suíça, Reino Unido e Itália (neste caso com a participação da biblioteca nacional de Florença e do ICCU - *Istituto Centrale per il Catalogo Unico*). A coordenação foi do próprio CENL.

Neste projecto foram definidos os pressupostos básicos da tecnologia de interoperabilidade a usar. Foi assim definido que a participação de cada biblioteca seria possível por uma de duas formas: através da disponibilização de um servidor Z39.50, ou de um servidor SRU - *Search and Retrieve via URL* (uma versão dos mesmo princípios do Z39.50, mas tecnicamente mais flexível e eficiente). Foi também acordada para este fim a definição do esquema de metadados "TEL Application Profile", baseado em elementos do dicionário Dublin Core.

O resultado desta iniciativa foi um portal de pesquisa bibliográfica, mantido de forma regular desde Março de 2005 (Figura 2). Na **Erro! A origem da referência não foi encontrada.** mostra-se um exemplo de uma sessão de pesquisa actual neste portal. Na primeira imagem pode-se ver do lado esquerdo uma lista com o total de registos encontrados em cada catálogo de cada biblioteca participante, e do lado direito da mesma imagem os resultados concretos encontrados no catálogo da BNP. Na imagem da direita vemos a expansão da informação para um registo da BND em concreto.

O portal actual incorpora já uma alteração significativa relativamente ao protótipo inicial. De facto, cedo se verificou a limitação do modelo de pesquisa distribuída em tempo real lançada em todos os catálogos, inerente ao uso das interfaces Z39.50 e SRU.

O primeiro problema é o da dificuldade em se garantir a resposta atempada de todos os servidores, ilustrado na **Erro! A origem da referência não foi encontrada.** pelos casos indicados com "Not responding". Este problema pode-se dever a excesso de carga no servidor em causa, a problemas no tráfego na Internet, ou simplesmente ao facto de o serviço estar inoperacional.

O segundo problema tem a ver com a impossibilidade de apresentar aos utilizadores os resultados segundo um critério único de relevância. O facto de cada servidor enviar as suas respostas em separado e ordenadas segundo os seus critérios próprios, faz com que não faça sentido ao portal considerar agregar e reordenar todos esses resultados. Em consequência, os resultados de cada catálogo são oferecidos em separado (sendo o que aparece no lado direito das imagens do exemplo).

Para contornar estes problemas concluiu-se que a melhor solução seria convencer as bibliotecas a concordar com a transferência dos seus metadados para o serviço central do portal TEL, onde já seria possível indexá-los. Os benéficos para os utilizadores seriam uma pesquisa mais rápida, e a possibilidade de ordenar os todos os resultados numa única lista.

Este novo modelo levou à adopção voluntária por algumas bibliotecas, já na fase derradeira do projecto TEL, do protocolo OAI-PMH, que dessa forma se prontificaram a transferir os seus registos para o serviço central. A BNP foi pioneira nesta experiência, cujos resultados foram importantes para convencer as mais renitentes da correcção desta decisão. Em concordância, esse modelo viria depois a ser generalizado no contexto do projecto TELplus.

Entretanto, após o projecto inicial que suportou o arranque do portal TEL, foi desenvolvido ainda entre 2005 e 2007 o projecto TEL-ME-MORE, que no essencial visou facilitar a adesão de mais 10 bibliotecas de países com adesão recente à Comunidade Europeia. Este projecto foi também financiado pelo programa IST – *Information Society and Technologies* da Comissão Europeia, mas agora já no âmbito do 5º Programa Quadro. Num outro passo, entre o final de 2006 e o final de 2008 desenvolveu-se o projecto EDL, liderado pela Biblioteca Nacional da Alemanha e que apoiou a adesão de mais nove bibliotecas (Bélgica, Grécia, Islândia, Irlanda, Liechtenstein, Luxemburgo, Noruega, Espanha e Suécia). Este projecto foi financiado pelo programa eContentPlus, e apoiou ainda o desenvolvimento das capacidades multilingue do portal. Finalmente, durante 2008 e 2009 decorreu ainda o projecto FUMAGABA, financiado pela Agência Suíça para o Desenvolvimento e Cooperação, que suportou a adesão de mais 8 bibliotecas de países não membros da Comunidade Europeia (Macedónia, Ucrânia, Moldávia, Albânia, Geórgia, Arménia, Bósnia Herzegovina e Azerbaijão). Entretanto outras bibliotecas aderiram individualmente com os seus meios próprios, tias como a Biblioteca Nacional da França.

PROJECTO TELPLUS

O projecto TELplus decorreu durante 2008 e 2009, com o objectivo de desenvolver novas soluções para alguns serviços da infra-estrutura TEL entretanto já obsoletos ou que nunca atingiram o seu potencial. Em concreto, foram desenvolvidos esforços pela BNP e IST em três frentes: foi redesenhada e desenvolvida uma nova infra-estrutura OAI-PMH; foram testadas novas soluções para suporte a serviços de pesquisa; e foram feitas experiências inovadoras sobre o uso do modelo FRBR para agregação de resultados de pesquisa.

REPOX – Infra-estrutura OAI-PMH

REPOX é uma plataforma para disponibilização de metadados por OAI-PMH ou também para recolha [3]. Inclui assim um servidor e um cliente, acrescidos de funções próprias para gestão de processos comuns em fornecedores de dados (*data providers*) e agregadores. Esta plataforma foi inicialmente desenvolvida pela BNP e INESC-ID, a pensar em necessidades internas da BNP e da rede PORBASE (FREIRE). Acabou no entanto por evoluir para uma plataforma genérica, adaptada durante o projecto TELplus às necessidades da rede TEL.

Na Figura 3 mostra-se um exemplo da interface de administração de uma instância dessa plataforma, na perspectiva do agregador. O caso refere-se a um teste com um grupo de bibliotecas que disponibilizam várias colecções de metadados para o portal TEL. Para os fornecedores de dados, a plataforma tem suporte para integração com servidores locais de bases de dados (permitindo assim ser acoplada a sistemas locais de gestão de bibliotecas) ou para importação directa de ficheiros. Esta plataforma está neste momento em fase de distribuição alargada pelos parceiros do TEL que não tenham tecnologia OAI-OMH própria. A sua aceitação tem sido extremamente positiva, o que levou a acordar adaptá-la também para a rede da Europeana.

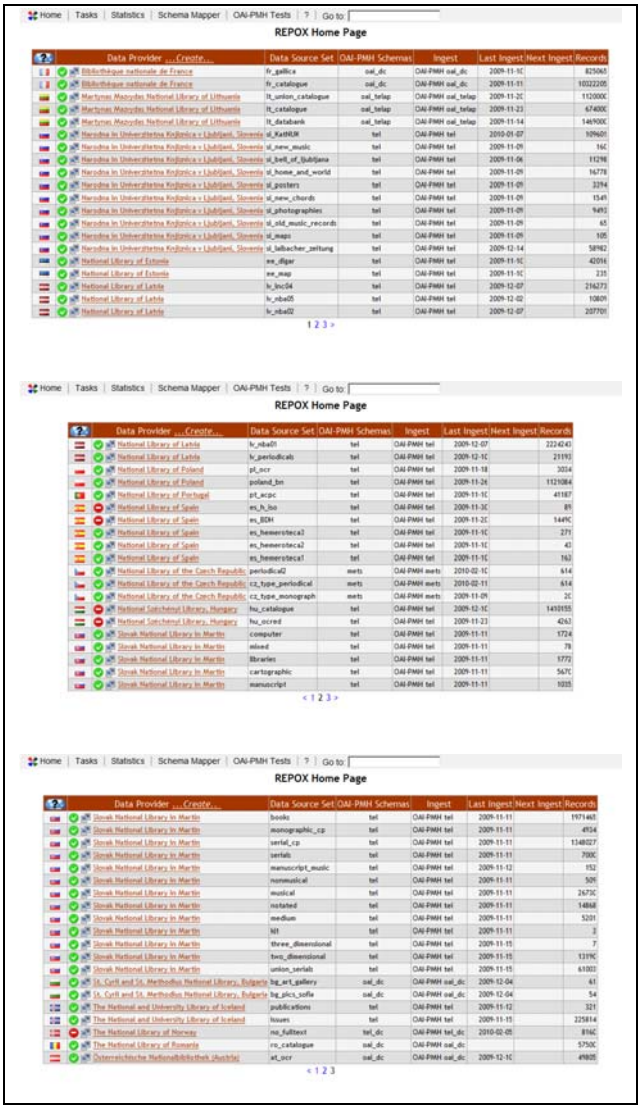

Figura 3: REPOX – A interface de agregador.

Indexação de texto

Outra tarefa relevante feita em parceria entre a BNP e o IST no projecto TELplus foi a indexação, pela primeira vez, de conteúdos em texto. Recordando o objectivo do portal TEL de oferecer um serviço unificado de pesquisa nos catálogos de todas as bibliotecas nacionais europeias, concluiu-se que seria uma grande melhoria ao serviço se fosse possível indexar juntamente com os registos bibliográficos também o texto das obras que existissem em formato digital (obras nascidas já digitais, ou entretanto digitalizadas).

Neste caso foi desenvolvida uma experiência de indexação de obras digitalizadas e submetidas a processamento por OCR (reconhecimento óptico de caracteres). O objectivo foi também o de testar a relevância para o TEL da tecnologia SOLR 9[4], beneficiando da experiência do INESC-ID na construção de índices deste género.

Fornecedor	Processo de Aquisição		Tamanho dos ficheiros SOLR (em Kbytes)
	Método de recolha	Total de páginas	
Austria	OAI-PMH	511.000	7.113.648
Czech Republic	OAI-PMH	2.579.511	4.727.796
Estonia	OAI-PMH	713.933	9.533.640
France	Manual	8.242.908	57.706.636
Hungary	OAI-PMH	237.914	2.094.880
Iceland	OAI-PMH	5.727.149	111.459.064
Latvia	OAI-PMH	195.075	6.165.564
Lithuania	HTML/HTTP	125.477	2.225.904
Norway	OAI-PMH	1.600.000	1.293.160
Poland	OAI-PMH	436.198	1.419.316
Slovakia	HTML/HTTP	185.000	0
Slovenia	OAI-PMH	328.502	5.890.612
Spain	OAI-PMH	3.033.525	10.275.680
Sweden	OAI-PMH	1.409.095	8.876.824
	Total	**25.325.287**	**228.782.724**

Tabela 1: Colecções em texto indexadas no projecto TELplus.

O paradigma FRBR na apresentação de resultados de pesquisa

O modelo conceptual FRBR veio apresentar um novo paradigma resultante do reconhecimento das limitações dos modelos MARC face ao potencial da realidade tecnológica emergente.

Neste projecto foi levada a cabo pela BNP e IST uma experiência para testar a relevância da aplicação desse modelo ara a apresentação dos resultados no portal TEL. Os registos utilizados, relatados na Tabela 2, estavam nos formatos MARC21 e UNIMARC (respectivamente 2/3 e 1/3), pelo que apenas se testou a relevância do modelo FRBR para apresentação de resultados de pesquisa, não tendo este estudo incidido sobre a catalogação.

Para o efeito foi desenvolvido um sistema que complementa um OPAC, recebendo deste uma lista de resultados de uma pesquisa e retornando esses resultados agregados segundo os princípios FRBR. A solução demonstrada pressupõe que o sistema FRBR tem acesso antecipado aos mesmos registos indexados no OPAC, os quais são pré-processados para o efeito e estruturados num esquema FRBR desenhado para este fim [5].

Na Figura 4 é apresentado um exemplo dos resultados dessa agregação por FRBR para o caso das obras de José Saramago. Na Figura 5 pode-se ver o ficheiro de autoridade criado automaticamente para o mesmo autor, apenas através do processamento automático dos registos bibliográficos.

O facto de o nome próprio do autor aparecer como "Zhoze" deve-se ao facto de terem sido usados nesta experiência apenas registos bibliográficos e não de autoridade, pelo que é mostrado por omissão o nome que é primeiro encontrado, o que depende da ordem com que os registos foram processados - neste caso parece que terá sido um registo de uma das bibliotecas de língua eslava; prevê-se corrigir isto no futuro, quando se processarem igualmente os registos de autoridade.

Tirando a nota anterior, pensamos serem objectivamente visíveis os resultados bastante encorajadores conseguidos com esta experiência. É importante realçar que tudo o que se apresenta foi conseguido apenas através de métodos de processamento automáticos, sem qualquer intervenção correctiva manual. Nesses métodos usou-se apenas o conhecimento das estruturas MARC e os conteúdos dos registos bibliográficos fornecidos (nesta experiência foi ainda possível detectar da parte de várias bibliotecas

Figura 4: Exemplo das obras de José Saramago, provenientes

uma gama alargada de práticas de catalogação não conformes com o esperado, o que lhes foi reportado).

Biblioteca	Registos Nobel
Biblioteca Nacional de Portugal	3.130
British Library	18.838
Royal Library of Belgium	1.084
National Library of Latvia	1.786
National Library of Lithuania	341
Russian State Library	20
National Library of the Czech Republic	2.945
Bibliothèque nacionale de France	28.302
German National Library	22.598
National Library of Spain	12.444
National Library of Servia	1.047
	92.535

Tabela 2: Colecções usadas na experiência FRBR no projecto TELplus.

Figura 5: Exemplo do ficheiro de autoridade criado automaticamente para "José Saramago"na experiência FRBR.

Em paralelo ao desenvolvimento do portal TEL nasceu a ideia do portal Europeana – The European Digital Library [6]. Este portal pretende ser um serviço para pesquisa de recursos em linha relativos à herança cultural e científica Europeia, alinhado com a chamada *Information Society i2010 Initiative* [7].

Em termos práticos pretende-se construir um agregador Europeu para os metadados de todas as bibliotecas, arquivos, museus, repositórios institucionais de universidades e centros de investigação, assim como de todas as demais entidades públicas ou privadas relevantes para o propósito. Os metadados devem ser recolhidos por OAI-PMH, e processados centralmente com técnicas semânticas sofisticadas que permitam não apenas suportar um serviço de pesquisa mas também serviços de navegação. Em última análise o portal poderá incluir funcionalidades de apoio ao acesso aos conteúdos, como resolução de identificadores, detecção de duplicados e mesmo suporte para comércio electrónico quando tal se justificar.

Esta iniciativa é coordenada pela EDL *Foundation* [8], sendo as actividades desenvolvidas na sequência de uma prova de conceito disponível em linha desde o final de 2008. Desde o primeiro momento que é objectivo desta iniciativa ser o mais inclusiva e abrangente possível, seja para os utilizadores (ver na Figura 6 o cuidado na interface multilingue) seja para os fornecedores de conteúdos (para o que tanto se pretende dar valor às grandes instituições patrimoniais nacionais como aos pequenos fornecedores locais).

Quanto aos conteúdos, pretende-se assim incluir conteúdos não só de texto e imagem (ver exemplos na Figura 7), mas também de vídeo e som.

A prioridade da iniciativa Europeana neste momento é para o desenvolvimento de serviços de recolha de metadados e de serviços de pesquisa sobre os mesmos. Para a recolha dos metadados estão em desenvolvimento vários canais utilizando a tecnologia OAI-PMH, inicialmente promovidos por redes sectoriais de instituições europeias (como por exemplo os casos da rede TEL das bibliotecas nacionais, ou a rede APEnet para o caso dos arquivos nacionais). Para este fim têm sido promovidos vários projectos, dois dos quais contam importante participação nacional: EuropeanaLocal e EuropeanaConnect.

PROJECTO EUROPEANA

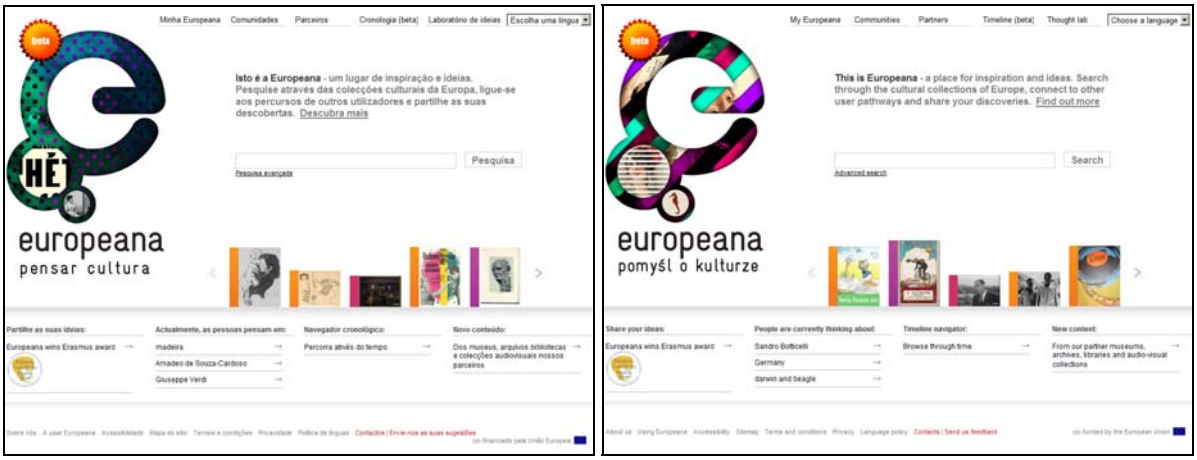

Figura 6: O portal EUROPEANA (evidenciando a sua interface multilingue)

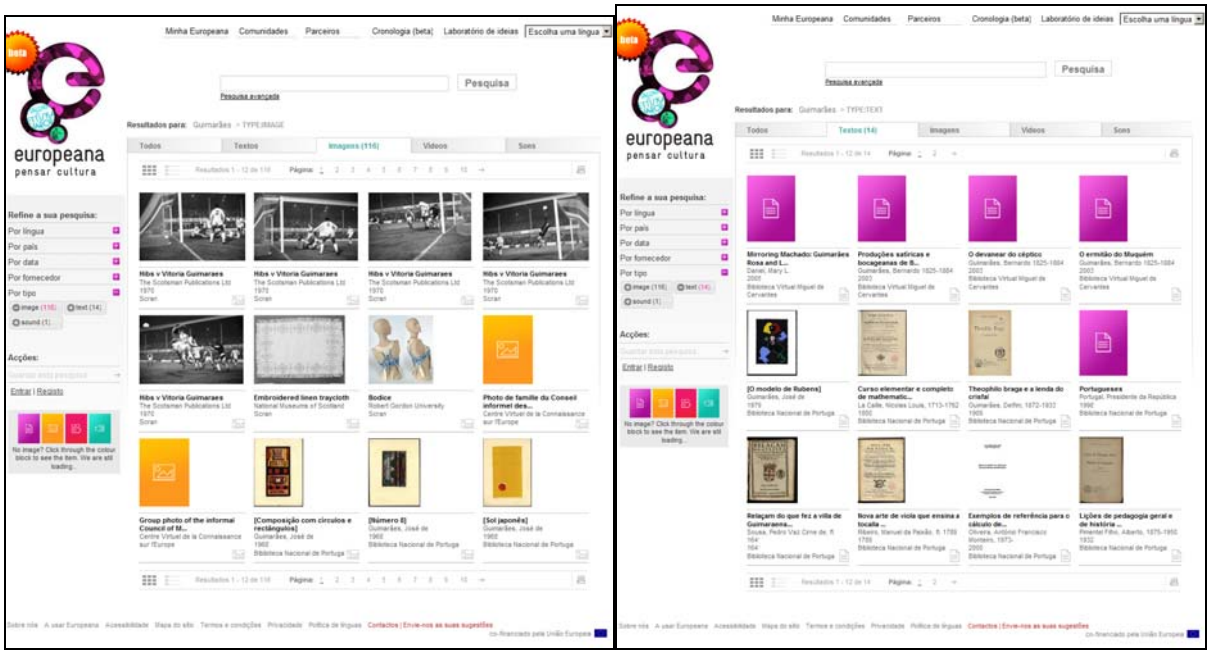

Figura 7: Resultados em texto e imagem de uma pesquisa por "Guimarães" no portal EUROPEANA.

Projecto EuropeanaLocal

O objectivo do projecto EuropeanaLocal é o de contribuir para a criação de novos canais de alimentação da Europeana, neste caso envolvendo repositórios regionais ou locais [9]. Para tal serão desenvolvidas soluções sob a mesma base tecnológica (OAI-PMH), mas levando agora em consideração os novos requisitos destes novos perfis de fornecedores de dados (geralmente organizações com recursos informáticos limitados, especialmente em termos de capacidade própria).

Este projecto iniciou-se em Junho de 2008, e irá decorrer por três anos. A representação nacional é assegurada pela FMNF - Fundação Museu Nacional Ferroviário, que tem como missão motivar e mobilizar os fornecedores de conteúdos nacionais de natureza local.

O INESC-ID apoia a FMNF neste objectivo, fornecendo apoio técnico às entidades interessadas e a tecnologia REPOX a quem o solicitar. No âmbito deste projecto irá ainda ser construído para o efeito um agregador nacional de metadados, cuja primeira versão pode ser vista na Figura 8.

Projecto EuropeanaConnect

O projecto EuropeanaConnect iniciou-se em Maio de 2009, e irá decorrer por 30 meses. O seu objectivo é o de contribuir para a construção da primeira versão verdadeiramente estável e funcional do serviço Europeana, com base na prova de conceito do protótipo actual. Mais uma vez a BNP e o IST são parceiros do consórcio. Para este projecto irão ser aproveitados resultados de projectos anteriores, tal como a plataforma REPOX já referida na descrição do TELplus. Mas irão ser também explorados resultados de outros projectos, especialmente do projecto DIGMAP.

Home | Tasks | Statistics | Schema Mapper | OAI-PMH Tests | ? | Go to:

REPOX Home Page

	Data Provider ...Create...	Data Source Set	OAI-PMH Schemas	Ingest	Last Ingest	Next Ingest	Records
	Arquivo Distrital do Porto	adporto	ese	Folder ese	2010-02-12		3074
	Biblioteca Digital Do Alentejo	bda	MarcXchange | ese	Folder ISO2709	2010-02-12		316
	Casa Fernando Pessoa	cfp	MarcXchange | ese	Folder MarcXchange	2010-02-12		1198
	CML - Departamento de Património Cultural	dpc	dpcxml | ese	Folder dpcxml	2010-02-12		98
	Câmara Municipal de Aveiro	aveiro	aveiroxml | ese	Folder aveiroxml	2010-02-12		999
	Fundação Museu Nacional Ferroviário	fmnf	fmnfxml | ese	Folder fmnfxml	2010-02-12		92
	Fundação Portuguesa das Comunicações	fpc	oai_dc | ese	Folder oai_dc	2010-02-12		2714
	Fundação Robinson	fr	frxml | ese	Folder frxml	2010-02-12		1015
	Hemeroteca Municipal de Lisboa	hemeroteca	MarcXchange | ese	Folder MarcXchange	2010-02-12		64
	Município de Vila Nova de Famalicão	famalicao	famalicaoxml | ese	Folder famalicaoxml	2010-02-12		513
	Museu da Região Flaviense	mrf	mrfxml | ese	Folder mrfxml	2010-02-12		341
	Museu dos Transportes e Comunicações	mtc	MarcXchange | ese	Folder MarcXchange			0

Figura 8: O agregador nacional de metadados do projecto EuropeanaLocal

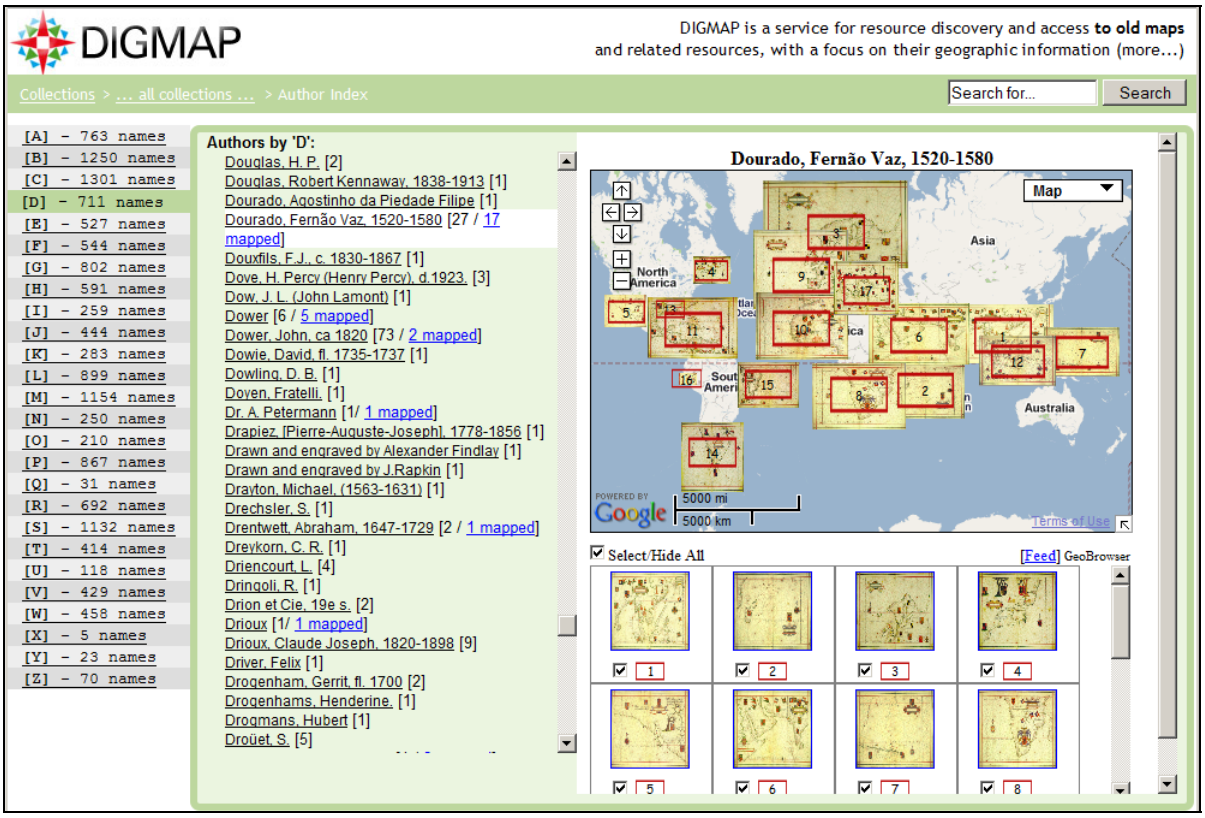

Figura 9: Exemplo de um índice de autores, com mapas indexados geograficamente.

PROJECTO DIGMAP

DIGMAP quer dizer "Descobrindo o Nosso Mundo Passado com Mapas Digitalizados". Este projecto decorreu de Outubro de 2006 a Setembro de 2008, tendo sido coordenado pelo IST e tendo tido oito parceiros, incluindo a BNP.

Neste projecto desenvolveram-se soluções para bibliotecas digitais geo-referenciadas, focadas especialmente em materiais históricos. Os resultados finais consistiram num grupo de serviços disponíveis na Internet [10] e em software de código livre para vários fins (este projecto constituiu mesmo o primeiro caso de aplicação da plataforma REPOX no exterior da BNP [11]). Na Figura 10 ilustra-se a página principal do portal demonstrativo.

Figura 10: Portal demonstrativo dos resultados do projecto DIGMAP.

Neste projecto utilizaram-se registos bibliográficos em MARC21 e UNIMARC de material cartográfico de quatro bibliotecas nacionais (Portugal, Bélgica, Estónia e Itália), assim como de outras fontes exteriores ao projecto (como por exemplo das bibliotecas nacionais da Austrália e Catalunha), grande parte deles com referências para versões digitalizadas de mapas antigos.

Com estes recursos demonstrou-se o valor em se enriquecerem metadados tradicionais, neste caso através de atributos geográficos. Para tal foi desenvolvido um *geoparser*, associado a um *gazetteer*, o que permitiu assim construir um serviço de enriquecimento dos registos. Esse serviço tem assim a capacidade de analisar os registos bibliográficos originais e detectar neles nomes de referências geográficas, o que permite acrescentar aos registos a respectiva informação de forma analítica.

Tipicamente, pode-se acrescentar automaticamente as coordenadas geográficas dos locais detectados, mas é também possível adicionar as coordenadas geográficas do rectângulo geográfico representado no mapa. Isto é possível de fazer automaticamente se estiver disponível nos metadados a escala do mapa e os metadados técnicos da imagem do mesmo (tamanho e resolução da imagem), ou pode-se fazer manualmente através de um editor concebido especialmente para esse efeito. Um exemplo da interface do serviço de indexação geográfica manual é mostrado na Figura 11.

Figura 11: Interface do serviço de indexação geográfica manual do projecto DIGMAP.

REALIZAÇÕES TECNOLÓGICAS

Nesta parte iremos sumarizar as principais realizações tecnológicas resultantes do envolvimento da BNP, IST e INESC-ID nos projectos e iniciativas referidas.

REPOX – Repositório de Metadados

A plataforma REPOX começou por ser idealizada para a Biblioteca Nacional Digital e para a PORBASE, mas o seu verdadeiro potencial acabou por ser concretizado no projecto TELplus e é neste momento uma aposta determinada no contexto da Europeana.

Esta é uma das contribuições nacionais mais relevantes dos parceiros nacionais para os serviços TEL e Europeana, pois constituirá a espinha dorsal da infra-estrutura OAI-PMH dessas redes. Esta tecnologia está já instalada em várias bibliotecas nacionais, indo agora ser alargada a outras entidades colaborando para a Europeana. Durante 2010 esperamos também vê-la instalada em fornecedores de conteúdos nacionais, no contexto do projecto EuropeanaLocal.

Na sua versão actual, a plataforma REPOX está vocacionada especialmente para metadados bibliográficos. No entanto o envolvimento da BNP e do IST no projecto EuropeanaConnect, e o caso nacional dado ao INESC-ID pelo projecto EuropeanaLocal, são motivações importantes para alargar as suas funcionalidades a outras áreas. O desafio aqui situa-se essencialmente nos processos de transformação dos esquemas de metadados, o que irá ser feito através de um serviço dedicado, o *Metadata Registry*.

Metadata Registry para a Europeana

No contexto do projecto EuropeanaConnect irá ainda ser desenvolvido um MDR - *Metadata Registry*. O conceito de MDR existe já definido na série de normas ISO/IEC 11179, mas irá ser utilizado por nós de forma mais alargada.

Desta forma iremos desenvolver uma infra-estrutura de gestão de esquemas de metadados da Europeana, onde serão registados e descritos de forma precisa todos os esquemas utilizados por todos os fornecedores de metadados. A forma como as descrições desses esquemas irão ser registadas irá seguir o estado da arte em tecnologia semântica, o deverá também permitir o mapeamento automático desses esquemas para o esquema da Europeana, o ESE - *Europeana Semantic Elements* [12].

Gazetteer e Geoparser

Dois resultados concretos do projecto DIGMAP foram os serviços de *gazetteer* (em termos simples, um thesaurus geográfico) e um *geoparser* (um serviço que identifica referências geográficas em textos, permitindo assim a geo-indexação dos mesmos).

Estes serviços irão ser agora reforçados e as suas novas versões definitivamente instaladas como parte da infra-estrutura da Europeana.

OUTRAS ACÇÕES RELEVANTES

Nesta parte focamos outras acções nacionais ou contando com a participação da BNP, INESC-ID ou IST e que poderão dar um importante contributo para o fornecimento de conteúdos para a Europeana.

Registo Nacional de Obras Digitalizadas

A nível nacional, é relevante a iniciativa da Biblioteca Nacional de Portugal que tem em vista a criação de um ponto central de registo de obras já digitalizadas, em processo de digitalização, ou que se pretendem digitalizar a curto prazo.

Actualmente, diversas instituições procedem à digitalização dos seus acervos de publicações, nomeadamente impressas. No entanto, não existem, actualmente, mecanismos de coordenação que permitam uma melhor rentabilização destas iniciativas. Este facto leva a que haja por vezes duplicação de esforços na digitalização de obras que integrem simultaneamente colecções de várias instituições interessadas na sua digitalização. Pode-se acrescentar ainda que a dispersão em termos das fontes e formas da disponibilização de obras digitalizadas, não potencia a sua melhor exploração por públicos alargados, limitando, assim, os efeitos práticos de cada iniciativa tomada individualmente. Esta realidade não favorece a melhor aplicação dos investimentos que estas iniciativas representam.

A constituição de um mecanismo a partir do qual se construa um ponto central de coordenação é, assim, uma necessidade a nível nacional. Esta urge para a implementação um serviço público que receba e disponibilize informação sobre a digitalização, de obras publicadas, por entidades portuguesas públicas ou privadas, e para efeitos de acesso público. Os objectivos desse serviço coadunam-se com as orientações que a Comissão Europeia tem vindo a promover desde os Princípios e Plano de Acção de Lund, agora especialmente reforçadas pela Europeana.

O controlo de publicações digitalizadas e disponíveis publicamente é entendido como extensão natural das funções de controlo bibliográfico nacional, dado respeitar a alternativas de acesso público a materiais que, no seu formato original de publicação, eram já eram abrangidos por essas funções. Pelas suas atribuições enquanto Agência Bibliográfica Nacional e entidade gestora da PORBASE – Catálogo Colectivo em linha das Bibliotecas Portuguesas, a Biblioteca Nacional proporciona o enquadramento técnico, funcional e organizacional adequado à implementação de um registo nacional de publicações digitalizadas.

O Registo Nacional de Obras Digitalizadas (RNOD) será um serviço em linha sobre digitalizações já efectuadas ou em intenção, com os respectivos dados

bibliográficos, endereços e condições de acesso. Incluirá também informação sobre as entidades responsáveis pela digitalização e acesso às obras.

No contexto da Europeana, o RNOD enquadra-se como agregador nacional de obras digitalizadas. Inicialmente, a Europeana foi alimentada pelas entidades que lhe disponibilizam as obras digitais directamente. No entanto, este serviço irá contar no futuro com recursos oriundos de milhares de entidades fornecedoras, o que levantaria obstáculos estruturais difíceis de gerir centralmente. Assim, no futuro a Europeana será alimentada a partir de entidades agregadoras, que funcionam como intermediários entre várias entidades que disponibilizam obras digitais e a Europeana.

Os agregadores podem estar organizados das mais variadas formas, como por exemplo em sectores particulares, por organização temática, ligadas a áreas geográficas, etc.

Perante a Europeana, os agregadores comprometem-se em fazer chegar os metadados das obras digitais, segundo as especificações técnicas publicadas. Desta forma um agregador oferece às respectivas entidades membros, serviços para transferência dos metadados, tanto entre elas, como entre o agregador e a Europeana. São também frequentes a disponibilização de serviços de processamento dos metadados, com vista a ao seu enriquecimento e conversão para o modelo de dados da Europena.

Pretende-se que o RNOD venha a oferecer também suporte estes serviços de agregação para a Europeana, potenciando ainda mais a sua melhor exploração das obras digitalizadas por públicos alargados.

EuDML

O projecto EuDML – *European Digital Mathematics Library*, iniciou-se a 1 de Fevereiro de 2010, com objectivos gerais semelhantes ao do DIGMAP. No concreto pretende-se no entanto agora desenvolver a infra-estrutura para a Biblioteca Digital Europeia de Matemática, promovida por vários editores e sociedades europeias de Matemática. No final esta biblioteca deverá vir a ter uma gama de serviços próprios inovadores, mas ainda assim deverá inter-operar com a Europeana.

A coordenação deste projecto cabe ao IST, e tecnicamente será alinhada com os requisitos da Europeana.

Aproveitando o contexto deste projecto, pretende-se ainda vir a criar, em colaboração com a Sociedade Portuguesa de Matemática e outras entidades do meio, a PtDML - a Biblioteca Nacional de Matemática.

CONCLUSÕES

O objectivo desta comunicação foi o de dar a conhecer os resultados do envolvimento a nível técnico do IST, INESC-ID e BNP, nas iniciativas internacionais que têm dado origem aos serviços TEL e Europeana. Pensamos ter ficado demonstrado o importante papel destas entidades, o qual até ao momento não tinha sido devidamente divulgado.

Estamos em crer que este tipo de envolvimento, quer a nível individual que como cooperação entre os parceiros nacionais, pode representar ainda um marco de referência para outros casos semelhantes. Por isso aqui fica também uma chamada de atenção às entidades nacionais relevantes, especialmente o Ministério da Cultura, o Ministério da Ciência e Ensino Superior, e ainda o Ministério da Educação

Em termos gerais, esta experiência pode servir também de referência para outros agentes no terreno, que assim podem ganhar uma nova confiança a nível nacional (em que se pode destacar a excelente oportunidade oferecida pelo projecto EuropeanaLocal).

NOTAS

[1] Imagem retirada de http://www.theeuropeanlibrary.org/portal/organisation/cooperation/cooperation_en.html

[2] http://search.theeuropeanlibrary.org

[3] http://repox.ist.utl.pt/

[4] http://lucene.apache.org/solr/

[5] Informação mais detalhada sobre esta experiência é fornecida na comunicação “O modelo FRBR e a descoberta de informação: a experiência do projecto TELplus”, igualmente apresentada neste congresso.

[6] http://www.europeana.eu

[7] http://europa.eu/i2010

[8] http://version1.europeana.eu/web/guest/edl-foundation/

[9] http://www.europeanalocal.eu/

[10] http://portal.digmap.eu

[11] http://bd1.inesc-id.pt:8080/repoxdigmap

[12] http://www.version1.europeana.eu/c/document_library/get_file?uuid=104614b7-1ef3-4313-9578-59da844e732f&groupId=10602

REFERÊNCIAS

FREIRE, Nuno; MAGUINHAS, Hugo; BORBINHA, José - Metadata Spaces: The Concept and a Case with REPOX – S. Sugimoto et al. (Eds.). ICADL 2006, LNCS 4312, p. 293-302, 2006. Springer Berlin / Heidelberg 2006, p. 293-302. ISBN 978-3-540-49375-4